DOCUMENT RESUME

ED 335 869 FL 019 322

AUTHOR Janopoulos, Michael
TITLE Using Written Recall Protocols To Measure Aspects of Holistic Assessment of Second Language Writing Proficiency: The Effect of Task Awareness.
PUB DATE 91
NOTE 57p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS College Faculty; *English (Second Language); Higher Education; Holistic Approach; Interrater Reliability; *Language Proficiency; Language Teachers; *Language Tests; Reading Comprehension; *Recall (Psychology); *Scoring; Second Language Instruction; *Writing Skills; Writing Tests

ABSTRACT

Holistic scoring is widely used to assess writing proficiency in English-as-a-Second-Language (ESL) composition. Written recall protocols have recently been used to investigate the relationship between how much holistic scorers comprehend of a given text and how high they rate the quality of that text. This study investigated whether readers approach a holistic evaluation task differently when they are focused (i.e., aware they will also be tested for comprehension) from when they are naive (i.e., unaware). Subjects were 12 college ESL composition teachers experienced in holistic assessment of writing proficiency. Each read, rated, and recalled one of two texts, then repeated the cycle with the other text. One text was superior to the other. Focused raters recalled significantly more overall but were unable to recall more of the better-written text. It is concluded that recall studies of the holistic reading process should use naive subjects to avoid the interference of an ensuing recall task. It is also suggested that insights into the holistic reading process can support, extend, and, to some extent, reconcile various general reading process/comprehension models. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED335869

Using Written Recall Protocols to Measure Aspects of Holistic Assessment of Second Language Writing Proficiency: The Effect of Task Awareness

Michael Janopoulos

University of Northern Iowa

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Janopoulos

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Running head: THE EFFECT OF TASK AWARENESS

BEST COPY AVAILABLE

Abstract

Holistic scoring is widely used to assess writing proficiency in both English and ESL composition. Written recall protocols have recently been employed to investigate the relationship between how much holistic scorers comprehend of a given text and how highly they rate the quality of that text. In this study, the author examined the issue of whether readers approach a holistic evaluation task differently when they are focused (i.e., aware they will also be tested for comprehension) from when they are naive (i.e., unaware). Subjects drawn from a pool of ESL composition teachers experienced in holistic assessment of writing proficiency read, rated, and recalled one of two texts. Then, they exchanged papers and repeated the read, rate, recall cycle. Analysis of the two sets of recall protocols reveals that, in contrast to their naive counterparts, focused raters recalled significantly more overall but were unable to recall more of the better-written text. The author concludes that recall studies of the holistic reading process should be done with naive subjects to avoid the interference caused by awareness of an ensuing recall

task and suggests that insights into the holistic reading process can help support, extend, and to some extent reconcile various general reading process/comprehension models.

Using Written Recall Protocols to Measure Aspects of Holistic Assessment of Second Language Writing Proficiency: The Effect of Task Awareness

Holistic scoring, which can be defined in generic terms as direct assessment of the overall quality of a writing sample (Davis, Scriven, & Thomas, 1981; White, 1985), is widely used in American universities to evaluate the writing proficiency of both native (NS) and non-native (NNS) speakers of English. Although there are several variations in the procedure (see Perkins, 1983; White, 1985), researchers in both English and ESL composition generally agree that holistic scoring is both a reliable (Perkins, 1983; White, 1985) and valid (Perkins, 1983) method for providing a direct assessment of writing proficiency. However, little has been done to investigate the holistic reading process, itself. Thus, questions concerning <u>how</u> holistic raters process text and <u>how much</u> they comprehend of what they have read persist among both researchers and practitioners in the fields of ESL and English composition.

One means of investigating the reading processes of holistic raters of writing proficiency is the written recall protocol. Recall protocols are a powerful tool in the study of prose comprehension. Their range of application includes analysis of text structure, the study of the organization of information in memory, and the study of how a text is processed from one task situation to another (Voss et al., 1982).

The widespread use of written recall protocols to measure reading comprehension supports Connor's (1984) assertion that recall is the operational definition of comprehension. Indeed, several researchers, including Meyer (1973a, 1973b, 1974, 1975, 1980a, 1980b, 1981a, 1981b, 1982), Meyer and McConkie (1973, 1974), Christopherson et al. (1981), Carrell (1983, 1984a, 1984b, 1986), Bernhardt (1983, 1984, 1985, 1986) and Lee (1986a, 1986b) have relied heavily on recall protocols to measure aspects of reading comprehension as they relate to the processing of texts by low-proficiency native speakers, ESL students at the university level, and foreign language learners at both the secondary and university level.

However, a potential drawback in using written recall protocols to explore the reading processes of holistic raters of second language (L2) writing proficiency lies in the question of whether the same group of subjects can be used in more than one recall protocol study, owing to the possibility that task awareness might influence (alter) the way in which they process NNS texts in an experimental environment. In other words, questions still persist as to whether holistic raters might somehow read "differently" or more "carefully" if they knew they were to be tested specifically on the content of what they had read. Thus, the issue of whether readers approach a holistic evaluation task differently when they are aware they will be tested for comprehension from when they are not forewarned is one that must be addressed in any research into the reading processes of holistic raters that may cause subjects to alter their normal reading strategies. Moreover, questions initially posed in the narrow range of holistic rating of writing proficiency may ultimately yield dividends along a much broader spectrum of reading research. Specifically, recall-protocol research into the reading processes of

holistic raters may provide insights into such aspects of the reading process, itself, as task awareness, depth of processing, and the organization of content in memory.

LITERATURE REVIEW

It is interesting to note that the overwhelming majority of research studies concerning text processing and the representation of meaning in memory have used subjects who were "focused" (i.e., who read a passage for the purpose of remembering it) rather than "naive" (i.e., who were unaware they were to be tested on the content of the passage). That is to say, regardless of whether the task involved writing a summary (e.g., Kintsch et al., 1977) or a recall protocol (e.g., Meyer, 1982), answering a set of multiple choice questions (e.g., Baker & Anderson, 1982), or holistically evaluating a composition and then rewriting it to more clearly reflect the author's intended meaning (e.g., Kaczmarek, 1980), the subjects were all aware that they would be tested in one way or another on what they had read.

That subjects in studies of this sort are focused is not surprising. Indeed, to attempt to execute a

reading experiment in which the subjects involved are not sensitized to the test environment would likely not be feasible owing to the extreme difficulty in controlling an experimental setting in which reactive factors might cause subjects to approach a reading task differently than usual, even if instructed to "read normally."

However, some studies touching upon the difference in reading performance of naive readers and of focused readers have shed some light on the problem of task awareness on reader response. These studies include those of Baker and Anderson (1982), Kaczmarek (1980), and Lee (1986a).

In a study designed to investigate the effects of inconsistent information on text processing, Baker and Anderson (1982) examined whether readers engaged in a focused reading task monitor their own comprehension in order to determine if ideas expressed in a text are consistent. They concluded that while subjects did engage in comprehension monitoring when reading logically inconsistent texts, they could not distinguish well between main ideas and details. The results obtained by Baker and Anderson also indicated

that specific instructions to read passages carefully in preparation for subsequent questions did not seem to influence how the subjects approached the reading task.

In their study, subjects were aware that they would be tested on what they had read. Moreover, some subjects were further focused in that they were instructed to be alert for inconsistencies in the texts they were to read. Nevertheless, Baker and Anderson reported that their highly-focused subjects spent no more time reading the target passages and were no more likely to notice textual inconsistencies during the reading process. Baker and Anderson were at somewhat of a loss to explain why this was so, although they did conjecture that "one possible reason for this lack of an effect of instructions is that the demand characteristics of the task were such that all subjects processed the text carefully in preparation for the test questions" (p.42).

Kaczmarek (1980) indirectly addressed the issue of focus in a study that compared teacher assessments of writing proficiency with those made by trained raters. Kaczmarek's study was designed to assess the validity and reliability of two different types of evaluation

methods, one subjective and the other objective. In her study, Kaczmarek used trained raters and writing instructors to evaluate essays written by international students in an American university.

In that phase of the study concerning holistic evaluation of student responses to a traditional essay question, writing instructors were asked to score essays on a six-point proficiency scale. Trained raters were also asked to read the essays, but their rating task was different. First, they were instructed to read each composition for an understanding of the writer's intended meaning. Then, they were asked to provide a holistic evaluation of the same essays, using a six-point scale that focused on comprehensibility rather than proficiency.

Kaczmarek was not looking specifically for the differences in recall between naive and focused raters of compositions written by NNS college students. Nevertheless, she was sufficiently struck by the results she obtained to acknowledge that the holistic scores of focused raters may well have been influenced by the fact that they were given a comprehension task in addition to a rating task.

Whether or not their scores were influenced by their being focused on the purpose of the reading task, however, is by no means clear, as Kaczmarek made no provisions in her experimental design for investigation of that specific question. Nor were her focused readers using the same criteria for holistic evaluation as her naive readers. In addition, as hers was a study of concurrent validity, her focused readers were trained holistic raters, but not experienced ESL composition teachers like her naive readers. Finally, Kaczmarek's procedures did not include having raters, either naive or focused, write recall protocols. Consequently, any conclusions drawn would be tenuous at best, especially insofar as rater comprehension, per se, was not an issue in her study.

In a few studies, readers were not oriented to read for the purpose of recalling the target passage (e.g., Carrell, 1983, 1984b). However, it is still not clear if holistic raters will, as a matter of course, read in the same way when reading specifically to form a general impression as subjects read when knowingly participating in a recall protocol study.

One study that did attempt to explore the question of how task awareness influences amount of recall was that of Lee (1986a). Lee used 320 NS American students enrolled in Spanish classes in two different universities. Consequently, variations in instructional procedures and student placement presented a potentially confounding source of variability. The same textbooks, however, were used in both programs at the four levels being tested.

All testing was carried out in the subjects' regular classrooms during regularly-scheduled times, although the experimenter was present during testing in only four of the 16 classes tested. All instructions were given in English, and the instructions contained in the test packets were varied so that an equal number of subjects were exposed to the respective test conditions.

All subjects were instructed to read a passage written in Spanish and then write a recall protocol of what they had read. Half of the subjects were required to produce protocols in Spanish, while the other half wrote in English. Half of each group were given prereading instructions alerting them to the fact that

they would be required to produce a protocol, while the other half received no prereading instructions beyond being told to read the passage at their own rate.

Protocols were analyzed on the basis of the type of idea unit matrix recommended by Bransford and Johnson (1972). Recall scores were analyzed by means of a three-factor analysis of variance. On the basis of this analysis, Lee concluded that "prereading directions alone are not sufficient to enhance recall of a passage" (p. 208).

It is tempting to generalize the results obtained by Lee to a setting in which trained and experienced raters of ESL writing proficiency are engaged in a holistic evaluation task. However, it must be remembered that Lee's subjects were university students performing under experimental conditions in which "subjects assume they will be performing some sort of task with the content of the passage" (p. 206).

Subjects' assumptions that the reading task involved attention to the content of the passage may or may not have been a confounding variable in Lee's experiment. However, in a task involving the assessment of writing proficiency, attention to content

becomes crucial to the issue of the role of comprehension in the process of holistic evaluation.

Specifically with regard to holistic evaluation, whether prior knowledge of comprehension testing significantly influences how closely one attends to detail in the target passage is a question that remains unanswered. In other words, will a subject read a passage differently (i.e., more carefully) for comprehension than he/she reads for the purpose of holistic evaluation? And will a subject, given such prior knowledge, still be able to attend to a holistic evaluation task while reading a text for comprehension?

THE STUDY

Purpose

The study reported here is the second of a two-part experiment in which subjects holistically rated the quality of texts generated by NNS college students and then wrote immediate recall protocols to test their comprehension of those texts (See Janopoulos, 1987, for details). The first part of the experiment revealed that holistic raters of L2 writing proficiency recalled significantly more of a better written text than of a text judged to be qualitatively inferior (see

Janopoulos, 1989 for details). The purpose of this, the second part of the experiment, was to determine if task awareness - specifically, prior knowledge that a holistic rating task was to be followed by a test of reading comprehension - affects the way in which holistic raters of L2 writing proficiency approach their reading task.

In this study, after the subjects had rated and recalled a NNS text as naive readers, they were instructed to rate and recall a second NNS text. As the subjects had just completed one "read/rate/recall" cycle, they were aware that they would be tested on their recall. Thus, in the experiment described here, the subjects were focused. Proceeding from the assumption that focused holistic raters will approach a reading task differently than their naive counterparts, this study posed the following two research questions:

1. Will focused holistic raters recall significantly more overall of two NNS texts than their naive counterparts, and;

2. Will focused holistic raters recall significantly more of a better written NNS text than of a less-well written NNS text?

Subjects

Twelve ESL composition instructors at the Ohio State University participated. Each of the 12 instructors had extensive prior experience in using holistic evaluation procedures to assess L2 writing proficiency.

Materials

Placement essays written by incoming NNS graduate students at Ohio State were used for all phases of the study. These essays, assigned as a one hour in-class writing task, were rated on a four-point scale that corresponds to the three proficiency levels used by Ohio State University's ESL Composition course sequence (106, 107, 108), plus Q (Qualified to bypass the ESL Composition course sequence).

Procedures

The Recall Task

In order to familiarize themselves with the rating standards, subjects were provided a list of holistic criteria and several model compositions one week prior to the experimental session. The session commenced with a discussion of the scoring criteria, followed by

a practice rating session during which results were discussed and subjects' scores were compared.

Once the researcher was satisfied that subjects were using the same general rating criteria, they were given a series of ten randomly selected compositions one at a time. After reading each text, subjects recorded their scores on individual tally sheets.

Two functions were served in this phase of the study. First, interrater reliability (Pearson r) was computed by comparing the scores assigned immediately after the original writing task with the scores assigned by each subject. Interrater reliability ranged from a low of .6877 to a high of .9129, with ten of twelve subjects achieving a Pearson r correlation coefficient in excess of .77.

Second, the consecutive scoring of several compositions created an environment which closely approximated that of a typical holistic rating session. In this way, subjects were conditioned to approach the rating task as naive readers, without forewarning of the subsequent recall task.

Once the appropriate holistic scoring environment was created, the next phase of the experiment began.

Subjects were given an eleventh essay to rate, but unlike the previous phase, they were not given the same essay. Rather, half received a level 107 (lower quality) text and half received a level 108 (higher quality) text (see Appendix A). As with all sample texts used in the experiment, prejudgments of these two texts were based upon initial raters' assessments and verified independently by the researcher and the director of the ESL Composition program.

Upon completing the holistic scoring task, subjects were directed to write a recall protocol of what they had read. Specific instructions followed guidelines established by Johnson (1970), with paraphrasing allowed when necessary (Frederiksen, 1977; Mandler, 1970). Subjects were permitted an unlimited time to write their protocols, but were not allowed to refer to the essays they had just read.

Next, subjects traded papers and repeated the process of reading and rating the sample texts (see Figure 1). Upon completion of this task, subjects were again instructed to write recall protocols of the compositions they had just read. Thus, subjects

approached the reading/rating task in a focused, or forewarned manner.

The Scoring Task

Generation of "Negotiated" Texts

Because the recall task in this study used texts written by NNS authors, steps had to be taken to insure that each instrument for scoring recall protocols accurately reflected the NNS author's intended meaning. Accordingly, three independent readers rewrote the texts to reflect what they felt was an accurate assessment of each NNS author's intended meaning. Rewritten texts were then compared in order to resolve differences in interpretation and reach consensus. These "negotiated" texts served as the basis upon which a scoring matrix for each of the two NNS compositions was produced (see Appendix B).

Generation of Weighted Propositional Counts

The two "negotiated" texts were analyzed using Meyer's (1973) proposition count procedure. The procedure, which assigns each idea unit (proposition) a value of relative importance to the text, is designed to extract a recall score from a written protocol.

Following this procedure, the researcher first compiled a list of all propositions contained in each text. Then, a colleague with experience in the procedure compared the lists to the texts and suggested modifications. Next, three other colleagues, using copies of the "negotiated" texts as guides, assigned each proposition a rating of from 1 (of negligible importance to the text) to 7 (extremely important to the text). The three ratings for each proposition were then averaged, with the resulting list of weighted propositional values creating a Protocol Scoring Matrix (PSM), or mean rating scale, for each text (see Appendix C).

Scoring of Written Recall Protocols

Subjects' written recall protocols were matched with the appropriate idea units to determine the number of correctly recalled propositions. The raw scores of all the idea units in each protocol were then added to compute the raw recall score for that particular protocol.

Scoring reliability for raw recall scores was then established by a second rater, who had been extensively trained in the procedure. Interrater reliability

(Pearson r) between researcher and independent rater was computed at .9695.

Data Analysis

Raw recall scores were converted into weighted recall scores by tallying the values assigned by the PSM to each correctly recalled proposition. These values were, in turn, converted into percentages by dividing them into the sum total of weighted idea units in each NNS-authored text. Thus, protocol scores based on the higher and lower quality NNS compositions were made amenable to comparison.

A two-way analysis of variance (ANOVA) was then computed to compare the mean recall scores of both groups of subjects (Focused and Naive) for both NNS texts (Lower and Higher). Followup procedures employed a one-way ANOVA.

Results

The null hypothesis for the first research question posed by this study predicted that there would be no significant difference between the total amount of content recalled by focused readers of two NNS texts and the total amount of content recalled by naive readers of the same texts.

Table 1 shows the means and standard deviations for holistic reader orientation (Focused versus Naive) and level of assessed proficiency (Lower versus Higher) of the two NNS texts read and recalled in this study. Examination of Table 1 reveals that focused readers produced a mean recall score of 52.07 (S.D.=13.82), while Naive readers produced a mean recall score of 35.72 (S.D.=15.16). The two-way Anova presented in Table 2 yielded an F-ratio that was statistically significant ($f(1,20)=8.56$, $p<.01$). On the basis of these results, the null hypothesis must be rejected. Thus, it can be concluded that Focused holistic readers as a group recalled significantly more of the two NNS texts than their Naive counterparts.

The null hypothesis for the second research question addressed in this study predicted that there would be no significant difference between the amount of content recalled by Focused readers of the lower-rated NNS text and the higher-rated NNS text. Table 1 shows that Focused readers of the Lower text produced a mean recall score of 53.63 (S.D.=14.07), while Focused readers of the Higher text produced a slightly lower mean recall score of 50.50 (S.D.=14.72). A one-way

ANOVA (see Table 3) yielded an F-ratio that was not statistically significant. On the basis of these results, the null hypothesis must be retained. In other words, Focused readers recalled both the higher-rated and lower-rated texts equally well.

Conclusions

Previous research (Janopoulos, 1987, 1989) has revealed that naive holistic raters of NNS writing proficiency recall significantly more of a better written text than of a text of lower quality. Thus, there appears to be a relationship between the quality of a NNS text and the amount of content a holistic rater of writing proficiency recalls of that text.

The focus of this study was on the influence of task awareness on readers engaged in a holistic scoring procedure involving NNS texts. The study posed two questions:

1. Would focused readers recall significantly more overall than naive readers? and :
2. Would focused readers be able to recall significantly more of a better written NNS text than of a NNS text of inferior quality?

Analysis of the data produced in this experiment indicate that task awareness alters the way in which a reader approaches a holistic scoring task. Specifically, results indicate that focused subjects recalled more of the texts they assessed for writing proficiency than did their naive counterparts, but they were unable to recall significantly more of the better written text. In other words, recall was enhanced, but the ability to make a qualitative (i.e., content recalled) distinction between NNS texts of differing qualities was lost.

Discussion

In a narrow sense, the intent of this study was to determine if focused holistic raters of NNS writing proficiency performed in a comparable manner to their naive counterparts with respect to (1) their overall comprehension of a pair of NNS texts, and (2) their ability to recall more of a better written NNS text. On the basis of the results reported above, it must be concluded that in studies of the reading processes of holistic raters of writing proficiency that use written recall protocols to measure reading comprehension, subjects should be naive; that is, unaware that they

will be tested on their recall of the content of the texts they have to read.

In a broader sense, however, the results of this study raise basic questions concerning the reading process, itself. Now that evidence has been provided that task awareness affects the way in which a holistic reader approaches a text, we need to investigate why that is so, what strategies are employed, and how those strategies work.

A wide variety of reading process/comprehension models have been advanced by researchers in the fields of cognitive psychology, information processing, and psycholinguistics (e.g., LaBerge & Samuels, 1974; Goodman, 1976; Kintsch & van Dijk, 1978; Stanovich, 1980; Rumelhart, 1980; Miller & Kintsch, 1980; Just & Carpenter, 1980; Bernhardt, 1986). In many ways, these models offer conflicting accounts of the reading process, but in some areas they are by no means contradictory. This is especially true in the area of comprehension monitoring, where, as Baker and Brown (1984) note, "comprehension monitoring activities are implicitly, if not explicitly, incorporated into several recent models of comprehension" (p.355).

Moreover, insights into the mechanisms of comprehension monitoring provided by research in one field can complement and extend the descriptive power of a reading process/comprehension model in another. An example of this can be seen in the applications of the construct of metacognition discussed by Baker, (1979) and Brown (1980) to the reading model of LaBerge and Samuels (1974).

Samuels and Kamil (1984) observe that the LaBerge and Samuels model, in both its early and revised forms, speaks of how a reader's limited attention capacity can by severely taxed by "the combined demands of decoding and comprehension," thereby resulting in what they term "attention switching"(p.197). Once the attention switching mechanism is activated, a shift in the allocation of processing capacity is initiated, and the skilled reader presumably applies any or all of what Brown (1980) identifies as the three main types of metacognitive skills - awareness, monitoring, and deployment of compensatory skills - to the reading task. Furthermore, it is plausible that attention switching can be triggered by other task demands than decoding and comprehending. For example, the attention

switching phenomenon posited by LaBerge and Samuels might also be seen as the trigger that activates the mechanism by which skilled readers employ different metacognitive strategies when reading for meaning (i.e., comprehension monitoring) than when reading for remembering (e.g., identifying important ideas and testing one's mastery of material).

In the present study, the fact that focused raters recalled significantly more overall but were unable to recall more of the better written of the two NNS texts may be explained in the following way: the multiple task demands imposed upon focused raters resulted in attention switching, which not only affected their allocation of processing capacity, but also influenced the selection of those metacognitive strategies and skills appropriate to both comprehension and recall.

The naive holistic raters in this study clearly read with comprehension, perhaps because when holistically assessing writing proficiency, readers cannot accomplish the latter without doing the former. To accomplish this, they likely employed the metacognitive strategy of comprehension monitoring. Focused raters, on the other hand, may well have

mobilized the metacognitive strategies associated with reading for remembering in addition to comprehension monitoring, which, in turn, resulted in better overall comprehension of both the lower and higher quality texts.

Yet their enhanced recall was apparently achieved at the expense of the ability to make significant qualitative distinctions between the two NNS texts, perhaps because a lesser amount of meaning needs to be accessed in order to activate the metacognitive skills needed to read for remembering. Moreover, perhaps in reading for remembering, comprehension becomes a means rather than an end in itself.

CONCLUSION

Clearly, the results reported in this study raise more questions than they answer, especially with regard to the reading strategies employed by holistic scorers of writing proficiency. Moreover, this study focused on an extremely specialized holistic reading task - native English speakers reading texts written by non-native English speakers - and considered only one variable in that task - prior knowledge of an ensuing test of recall. Further research that investigates the

holistic reading process, itself, more directly (using, for example, eye-tracking technology or complete propositional analyses of recall protocols) is also needed to provide a clearer picture of the influence of task demands on not only holistic reading for the purpose of assessing writing proficiency, but all other types of reading activities as well.

References

Baker, L. (1979). Do I understand or do I not understand: That is the question. (Reading Education Report no. 10). Urbana: University of Illinois, Center for the Study of Reading. (ERIC Document Reproduction Report No. ED 174 948).

Baker, L. & Anderson, R.I. (1982). Effects of inconsistent information on text processing: Evidence for comprehension monitoring. Reading Research Quarterly, 17(2), 281-293.

Baker, L. & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Pearson et al. (Eds.), Handbook of Reading Research. New York: Longman.

Bernhardt, E.B. (1983). Three approaches to reading comprehension in German. Modern Language Journal, 67(2), 111-115.

Bernhardt, E.B. (1984). Toward an information processing perspective in foreign language reading. Modern Language Journal, 68(4), 322-331.

Bernhardt, E.B. (1985). Cognitive processes in L2: An examination of reading behaviors. In J. Lantolf & A. LaBarca (Eds.), Delaware Symposium on Language

Studies Research on Second Language Acquisition in Classroom Setting. Norwood, NJ: Ablex.

Bernhardt, E.B. (1986). A model of L2 text reconstruction: The recall of literary text by learners of German. In A. LaBarca (Ed.), Issues in L2: Theory as Practice/Practice as Theory. Norwood, NJ: Ablex.

Bransford, J. & Johnson, M. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. Journal of Verbal Learning and Verbal Behavior, 11, 717-726.

Brown, A. L. (1980). Metacognitive development and reading. In R. J. Spiro et al. (Eds), Theoretical Issues in Reading Comprehension. Hillsdale, NJ: Erlbaum.

Carrell, P.L. (1983). Three components of background knowledge in reading comprehension. Language Learning, 33, 183-207.

Carrell, P. L. (1984a). Text as interaction: Some implications of text analysis and reading research for ESL composition. Paper presented at the 18th Annual Meeting of TESOL, Houston, Texas.

Carrell, P.L. (1984b). Evidence of formal schemata in second language comprehension. Language Learning, 34, 87-113.

Carrell, P.L. (1984c). The effects of rhetorical organization on ESL readers. TESOL Quarterly, 18(3), 441-469.

Carrell, P.L. (1986). Interaction of content and formal schemata in ESL reading. Paper presented at the 20th Annual Meeting of TESOL, Anaheim, CA.

Christopherson, S. et al. (1981). The effect of two contextual conditions on recall of a reading passage and on thought processes in reading. Journal of Reading, April, 573-578.

Davis, B., Scriven, M., & Thomas, S. (1981). The Evaluation of Composition Instruction. Inverness, CA: Edgepress.

Frederiksen, C.H. (1975). Representing logical and semantic structure of knowledge acquired from discourse. Cognitive Psychology, 7, 371-458.

Goodman, K.S. (1976). Reading: A psycholinguistic guessing game. In H. Singer & R. Ruddell (Eds.), Theoretical Models and Processes of Reading, (2nd

ed.). Newark, DEL: International Reading Association.

Janopoulos, M. (1987). The role of comprehension in holistic evaluation of second language writing proficiency at the university level. Unpublished doctoral dissertation, The Ohio State University, Columbus, OH.

Janopoulos, M. (1989). Reader comprehension and holistic assessment of second language writing proficiency. Written Communication, 6(2), 218-237.

Johnson, R.E. (1970). Recall of prose as a function of the structural importance of the linguistic units. Journal of Verbal Learning and Verbal Behavior, 9, 12-20.

Just, M.A. & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. Psychological Review, 87, 329-354.

Kaczmarek, C.M. (1980). Scoring and rating an essay task. In J.W. Oller & K. Perkins (Eds.), Research in Language Testing. Rowley, MA: Newbury House.

Kintsch, W. et al. (1977). Summarizing scrambled stories. Memory and Cognition, 5, 547-552.

Kintsch, W. & Van Dijk, T. (1978). Toward a model of discourse comprehension and production. Psychological Review, 85(1), 363-394.

LaBerge, D. & Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. Cognitive Psychology, 6, 293-323.

Lee, J.F. (1986a). On the use of the recall task to measure L2 reading comprehension. Studies in Second Language Acquisition, Vol. #9, June. New York: Cambridge University Press.

Lee, J.F. (1986b). The effects of three components of background knowledge on second language reading. Modern Language Journal, 70(49), 350-354.

Mandler, J. M. (1978). The code in the node: The use of a story schema retrieval. Discourse Processes, 1(1), 14-35.

Meyer, B.J.F. (1973a). Identifying variables in prose. (ERIC Document Reproduction Service No. ED 082 218)

Meyer, B.J.F. (1973b). Structure in prose: Identification and effects on memory. (ERIC Document Reproduction Service No. ED 076 979)

Meyer, B.J.F. (1974). Effect of position of information in a passage's organizational structure

of recall. (ERIC Document Reproduction Service No. ED 090 525)

Meyer, B.J.F. (1975). The Organization of Prose and Its Effects on Memory. Amsterdam: North Holland.

Meyer, B.J.F. (1980a). Text structure and its use in the study of reading comprehension across the adult life-span. (ERIC Document Reproduction Service No. 184 110)

Meyer, B.J.F. (1980b). Signaling in text and its interaction with reader strategies. (ERIC Document Reproduction Service No. ED 185 551)

Meyer, B.J.F. (1981). Prose analysis: Purposes, procedures, and problems. (ERIC Document Reproduction Service No. ED 201 972)

Meyer, B.J.F. (1982). Reading research and the composition teacher: The importance of plans. College Composition and Communication. 33(1), 37-49.

Meyer, B.J.F. & McConkie, G.W. (1984). Effect of position of information in a passage's organizational structure on recall. (ERIC Document Reproduction Service No. ED 090 525)

Miller, J.R. & Kintsch, W. (1980). Readability and recall of short prose passages: A theoretical analysis. Journal of Experimental Psychology: Human Learning and Memory, 6, 335-354.

Perkins, K. (1983). On the use of composition scoring techniques, objective measures, and objective tests to evaluate ESL writing ability. TESOL Quarterly, 17(4), 651-671.

Rumelhart, D.E. (1980). Schemata: The building blocks of cognition. In R.J. Spiro et al. (Eds.), Theoretical Issues in Reading Comprehension. Hillsdale, NJ: Erlbaum.

Samuels, S.J. & Kamil, M.L. (1984). Models of the reading process. In P.D. Pearson et al. (Eds.), Handbook of Reading Research. New York: Longman.

Stanovich, K.E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. Reading Research Quarterly, 16, 32-71.

Voss, J.F. et al. (1982). Prose comprehension and memory. In R.C. Puff (Ed.), Handbook of Research Methods in Human Memory and Cognition. New York: Academic Press.

White, E.M. (1985). Teaching and Assessing Writing. San Francisco, CA: Jossey-Boss.

Table 1

Means and Standard Deviations of Recall Scores by Reader Orientation and Proficiency Assessment

Reader Orientation	Assessed Proficiency								
	Lower			Higher			Overall		
	n	$\bar{x}$	$\hat{\sigma}$	n	$\bar{x}$	$\hat{\sigma}$	n	$\bar{x}$	$\hat{\sigma}$
Naive	6	27.28	9.09	6	44.17	15.86	12	35.72	15.16
Focused	6	53.63	14.07	6	50.50	14.72	12	52.07	13.82
Overall	12	40.45	17.80	12	47.34	14.96	24	43.89	16.46

Table 2

Analysis of Variance of Reading Comprehension by Reader Orientation and Proficiency Assessment

Source	df	SS	MS	F	p
Proficiency Assessment (A)	1	284.1440	284.1440	1.52	0.2324
Reader Orientation (B)	1	1602.9541	1602.9541	8.56	0.0084*
Proficiency by Orientation (AB)	1	601.6010	601.6010	3.21	0.0883
Error (S/AB)	20	3746.8225	187.3411		
Total	23	6235.5218			

*$p < .01$

Table 3

Analysis of Variance of Reading Comprehension (Focused/Higher by Focused/Lower)

Source	df	SS	MS	F	p
Model (A_1B_2/A_2B_2)	1	29.4220	29.4220	0.14	.7143
Error	10	2074.3450	207.4345		
Total	11	2103.7670			

Subjects	Reader Orientation*	
	Naive	Focused
1	L	H
2	H	L
3	L	H
4	H	L
5	L	H
6	H	L
7	L	H
8	H	L
9	L	H
10	H	L
11	L	H
12	H	L

* L = Lower-rated (107) text

H = Higher-rated (108) text

Figure 1.

Order of administration of Naive and Focused reading tasks with respect to target texts

Appendix A

Original Texts (Levels 107 & 108)

Composition #1, Level 107

For giving an answer to the question if it's better to learn to be competitive or cooperative it is nessessary to consider a few faktors which are important

- In which society do the people live and how it is organized
- the personelity of the individual and the individual abilities.

Everyone will agree with the statement that it is important to learn both of these abilities, to be competitive and cooperative.

And everyone will agree that every society has developed its own way to selekt people at school in business everywhere.

Ther are differences between the social aims with an influence on the individual aims. In America the organization of the society is much more competitive oriented as in some states of the eastern hemisphere. And as far as I know are the people in Japan much more oreintated in a cooperative style of working and living but they are very competitive as a nation or as a company - as every kind of a group.

The reasons for these different kinds of social orgainzation can be found in the historical development and the conditions of living.

Japan, for many years isolated from any influence and with little space for too many people was forced to develop a very cooperative style of living. The development of the USA was determind of individual activities. So they developed a more competitive style of social orgainization.

It is important to see these historical and social criterias before answering the question if its better to learn a competitive or cooperative style of living. No individual person can be seen isolated from the society to which it belongs.

Composition #2, Level 108

I think this is a rather controversial problem. It is difficult for me to say that one is definitely better than the other since it may depend on the importance of the matter of the situation.

There are times when taking a quick and decisive action is more desirable. For instance, you got a very serious disease and your doctor told you that you need an operation which is very dangerous. You were also told that you might die if you don't have the operation. In such a case, it would be very hard to make a quick decision, but you should do so since the matter is very urgent. For another example, I go to a restaurant with some of my friends and we are supposed to have only 30 minutes in that restaurant because we are going to a movie afterwards. In that particular situation, I have to make a quick decision for my own sake. Otherwise, I may not be able to finish my meal or miss some part of the movie.

There are many other cases when it is much better to think something over carefully, rather than to make a hasty choice. It is more advisable to be careful especially when you have to make a very important decision in your life. For example, when you decide what you'll study in college, you should take time and think over and over since your major will be related to your career.

For another example, I can say when you choose your spouse. That kind of decision is what you have to live with it all your life and what you can not change easily even though you want to.

In conclusion, I would say that being carefull and taking enough time is more preferrable if the issue is not very urgent and doesn't need to be done right away because you'll probably make less mistakes and repent less if you do so.

Appendix B

Negotiated Texts

COMPOSITION #1 (LEVEL 10?

In order to give an answer to the question of whether it's better to learn to be competitive or cooperative it is necessary to consider a few factors which are important:

- In which society do the people live and how it is organized;
- The personality of the individual and the individual abilities.

Everyone will agree with the statement that it is important to learn both of these abilities, to be competitive and cooperative.

And everyone will agree that every society has developed its own way to select people at school and in business everywhere.

There are differences between the social aims that have an influence on the individual aims. In America the organization of the society is much more competitively oriented than in some states in the eastern hemisphere. As far as a know the people in Japan are much more oriented in a cooperative style of working and living but they are very competitive as a nation or as a company - whenever they function as a group.

The reasons for these different kinds of social organization can be found in the historical development and the conditions of living.

Japan, isolated for many years from any influence and with little space for too many people, was forced to develop a very cooperative style of living. In contrast, the development of the USA was determined by individual activities. So Americans developed a more competitive style of social organization.

It is important to see these historical and social criteria before answering the question of whether it's better to learn a competitive or cooperative style of living. No individual person can be seen as isolated from the society to which he/she belongs.

COMPOSITION #2 (LEVEL 108

I think the question of quick action versus deliberation is a rather controversial problem. It is difficult for me to say one is definitely better than the other since the choice between alternatives may depend on the importance of the matter of the situation.

There are times when taking quick and decisive action is more desirable. For instance, you got a very serious disease and your doctor told you that you need an operation which is very dangerous. You were also told that you might die if you didn't have the operation. In such a case, it would be very hard to make a quick decision, but you should do so since the matter is very serious. For another example, I go to a restaurant with some of my friends and we are supposed to have only 30 minutes in that restaurant because we are going toa movie afterwards. In that particular situation, I have to make a quick decision about what to order for my own sake. Otherwise, I may not be able to finish my meal or may miss some part of the movie.

There are many other cases when it is much better to think something over carefully, rather than to make a hasty choice. It is more advisable to be careful, especially when you have to make a very important decision in your life. For example, when you decide what you'll study in college, you should take time and think it over and over since your major will be related to your career.

For another example, I can say when you choose your spouse. That kind of decision is what you have to live with all your life and what you can not change even though you want to.

In conclusion, I would say that being careful and taking enough time is preferable if the issue is not very urgent and action doesn't need to be taken right away because you'll probably make fewer mistakes and repent less if you do so.

Appendix C

Protocol Scoring Matrices for Higher (FORM A) and Lower (FORM B) Texts

FORM A, Text #2

	WEIGHTED SCORES	IDEA UNITS
1.	3.33	I think (OPINION)
2.	5	" " the question of quick action versus deliberation is a problem.
3.	4	The problem is rather controversial.
4.	3.33	It is hard to say one is definitely better than the other...
5.	3.67	Since (Because)
6.	6.33	The choice between alternatives may depend on the importance of the situation.
7.	6	Sometimes taking quick and decisive action is more desirable.
8.	4	For instance (EXAMPLE #1)
9.	5.67	You have a disease.
10.	5.33	The disease is very serious.
11.	4.33	Your doctor told you.
12.	5.33	" " told you THAT YOU NEED AN OPERATION.
13.	5.33	The operation is dangerous.
14.	4.67	(BUT)
15.	6	... You may die without it.
16.	5.33	In this case, it would be very hard to make a quick decision.
17.	3.67	But

18.	4.33	You should do so...
19.	4	Since... (Because)
20.	4.67	... the matter is very urgent.
21.	4.33	For another example (EXAMPLE #2).
22.	5	I (the author) go to a restaurant...
23.	3.33	... with some of my friends...
24.	6.33	... and we only have 20 minutes to eat
25.	4.67	Because...
26.	5	... we are going to a movie afterwards.
27.	6	In this situat on, I have to make a quick decision.
28.	3.67	The decision is about what to order.
29.	2.33	The decision is FOR MY OWN SAKE.
30.	4	Otherwise (if I don't make a quick decision)...
31.	3.67	I may not be able to finish my meal...
32.	3	... or
33.	4	... I may miss part of the movie.
34.	3.33	(ON THE OTHER HAND)
35.	5.67	There are cases when it is much better to think something over carefully...
36.	4.33	" " carefully RATHER THAN
37.	4.67	... making a hasty choice.
38.	5.33	It is more advisable to be careful...
39.	6.33	" " careful WHEN YOU HAVE TO MAKE A VERY IMPORTANT DECISION IN YOUR LIFE.
40.	4.33	For example (EXAMPLE #1)

41. 4 — You should take your time to think it over...

42. 4.67 — " " think it over WHEN YOU DECIDE

43. 5 — " " decide WHAT YOU'LL STUDY IN COLLEGE...

44. 3.67 — Since (Because)

45. 4 — Your major will be related to your career.

46. 4.33 — For another example (EXAMPLE #2)

47. 3.33 — Take your time...

48. 5.67 — ... when you choose your spouse.

49. 5 — You have to live with that kind of decision...

50. 5.67 — " " decision ALL YOUR LIFE...

51. 5 — ... and you cannot change (that decision)...

52. 4 — ... even though (if) you want to.

53. 3.67 — In conclusion...

54. 4.33 — ... being careful

55. 5.67 — ... and taking enough time

56. 6 — ... time IS PREFERABLE

57. 4 — ... if

58. 5 — ... the issue is not very urgent

59. 4.33 — " " urgent AND (action) DOESN'T NEED TO BE DONE (taken)

60. 4.67 — ... right away (immediately)

61. 4 — ... because

62. 5 — ... you'll probably make fewer mistakes

63. 4 " " make fewer mistakes AND REPENT LESS

64. 3.33 ... if

65. 3.67 if YOU DO SO (take your time).

TOTAL

295.65

FORM B, Text #1

	WEIGHTED SCORES	IDEA UNITS
1.	4.67	It is necessary to consider a few factors.
2.	4.33	" " a few factors WHICH ARE IMPORTANT.
3.	3	In order to give an answer to the question...
4.	4.67	" " of whether it is better to learn to be competitive...
5.	4.67	... or cooperative.
6.	6	In which society do the people live?
7.	6	How is it organized?
8.	3.67	(AND)
9.	6.33	The personality of the individual.
10.	5.67	Individual abilities.
11.	2.33	Everyone will agree with the statement...
12.	4.33	" " the statement THAT IT IS IMPORTANT TO LEARN BOTH OF THESE ABILITIES (cooperation and competitiveness).
13.	2	Everyone will agree...
14.	3.67	" " agree THAT EVERY SOCIETY HAS DEVELOPED ITS OWN WAY...
15.	3.33	" " its own way TO SELECT PEOPLE...
16.	3	" " to select people AT SCHOOL...
17.	3	" " IN BUSINESS...
18.	2.33	" " EVERYWHERE.
19.	5.33	There are differences between the social aims...

20.	5.33	" " aims WHICH INFLUENCE INDIVIDUAL AIMS.
21.	5.67	In America...
22.	6	organization of society is much more competitively oriented...
23.	4.67	" " oriented THAN IN SOME STATES IN THE EASTERN HEMISPHERE.
24.	1.33	As far as I know...
25.	6	In Japan...
26.	6.33	" " the people are much more oriented in a cooperative style of working...
27.	5.33	" " of working AND LIVING,
28.	4.67	but
29.	6.33	they are very competitive as a nation
30.	3.33	or
31.	5.33	as a company -
32.	6	whenever they function as a group.
33.	4.67	There are reasons for these different kinds of social organizations.
34.	5	These reasons can be found in the historical development...
35.	5.67	" " and in the conditions of living.
36.	5.67	Japan has been isolated from outside influence...
37.	4	" " influence FOR MANY YEARS...
38.	5.67	" " and has had little space for too many people...
39.	3.67	" " people FOR MANY YEARS.
40.	4.67	(CONSEQUENTLY)

41.	6	Japan was forced to develop a very cooperative style of living.
42.	4.67	In contrast...
43.	5.67	The development of the USA was determined by individual activities.
44.	5	So (consequently)
45.	6	Americans developed a more competitive style of social organization.
46.	4.33	It is important to see these historical and social criteria...
47.	4	" " criteria BEFORE ANSWERING THE QUESTION OF WHETHER IT IS BETTER TO LEARN TO BE COMPETITIVE OR COOPERATIVE.
48.	4	No individual person can be seen as isolated from...
49.	3.67	" " from THE SOCIETY...
50.	2.33	" " society WHICH HE/SHE BELONGS TO.

TOTAL

229.34